



# Search for $H$ to $b\bar{b}$ in association with single top quarks as a test of Higgs boson couplings

Simon Fink

## Abstract

The production of a Higgs boson in association with a single top quark is one of a few channels which are sensitive for not yet excluded anomalous couplings of the Higgs boson to fermions. Multivariate analysis tools are used for the reconstruction and classification of signal events, where the Higgs boson decays into bottom quarks and the single top quark decays leptonically. In this conference report the most recent results at time of the conference using the full dataset recorded by the CMS detector at 8 TeV are presented.

**Keywords:** Single Top, Higgs, Yukawa coupling, Event reconstruction, MVA

## 1. Introduction

Since the discovery of a new particle by the CMS and ATLAS collaborations in 2012 [1, 2], all measurements are found to be consistent with the Higgs Boson predicted by the standard model (SM). In order to find any deviations from the SM and therefore possible signs of new physics beyond the SM, the properties of the particle have to be measured with ever increasing precision. One of these properties is the value of  $y_t$ , the Yukawa coupling of the Higgs boson to top quarks. Most measurements probing this property are only sensitive to the absolute value of  $y_t$ , rather than its sign, as quantities such as the rate of Higgs boson production in association with top-quark pairs depend only on  $|y_t|^2$ . The cross section for the production of Higgs bosons in association with single top quarks is particularly sensitive to the sign of  $y_t$ , or, more precisely, the relative phase of the couplings of the Higgs to the top quark and to W bosons. The two amplitudes lead to a nearly total destructive interference, with an expected cross section of 18.3 fb at 8 TeV in the standard model [3]. Any alteration of the relative phase of the two amplitudes would result in an enhancement of the cross section and this process is therefore very sensitive to enhancements in models beyond the standard model.

## 2. Analysis overview

The analysis which is described here focusses on the search for a signal process under the hypothesis of  $y_t = -1$ . Due to constructive interference the production cross section of this process would be enhanced by a factor of  $\sim 13$  up to  $\sigma = 234$  fb.

### 2.1. Event Topology and Event Selection

The final topology for the searched process consists of one isolated lepton, missing transverse energy, four bottom quarks as well as one distinct light forward jet. The Feynman graph for this process can be seen in Figure 1. In order to enhance the signal fraction in all considered events we apply a selection where one tightly isolated lepton, either an electron or a muon, is required; the missing transverse energy due to the escaping neutrino has to be larger than 45 GeV in the electron channel and larger than 35 GeV in the muon channel and jets have to have a transverse momentum greater than 20 GeV, if they lie in the central part of the detector ( $|\eta| < 2.4$ ), and greater than 40 GeV, if they lie in the outer detector regions ( $|\eta| > 2.4$ ). One of these jets has to be untagged, and at least 4 jets have to have a transverse momentum greater than 30 GeV. The remaining events are sorted according to the multiplicity of b-tagged jets into two

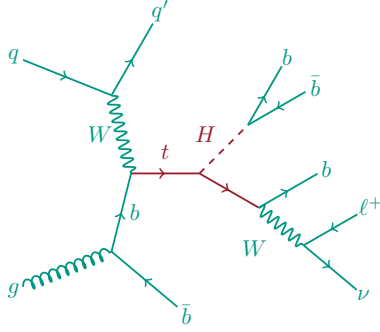


Figure 1: A characteristic Feynman diagram for the final state studied in this analysis.

independent signal regions and a control regions, events with three b-tagged jets form the three tag (3T) signal region, events with four b-tagged jets form the four tag (4T) region and events with two tags are sorted into the control region (2T) which is predominantly populated by  $t\bar{t}$  events.

## 2.2. Top pair production modelling

As mentioned before top pair production is the dominant background in this analysis and two approaches are used to model it. One approach is relying completely on the provided Monte Carlo simulation of this process, the alternative approach is using a data-driven method trying to model the process in our signal region based on the kinematics seen in the control region.

Events from the two tag control region are given a weight according to the probability of an event with the same kinematics to appear the 3T or 4T region. Consider an event with  $n$  reconstructed jets in which each jet has a unique b-tagging efficiency,  $\epsilon(p_i, f_i)$ , that is dependent on the transverse momentum and flavor of the jet. The probability of a selected event (but for the number of b tags) to have  $m$  tagged jets is

$$P_m = \sum_{comb} \prod_{i=1}^m \epsilon(p_i, f_i) \cdot \prod_{j=m+1}^n (1 - \epsilon(p_j, f_j)). \quad (1)$$

The final weight for each event is then calculated in the following way:

$$w = P_3/P_2 \quad \text{or} \quad P_4/P_2 \quad (2)$$

## 2.3. Event Reconstruction

The small signal-to-background ratio after the selection makes it essential to exploit multivariate techniques to discriminate signal processes from background processes. In order to construct observables which have a

good separation power, we employ two dedicated "jet-assignment MVAs", where the jets in an event are assigned to the quarks in the final state of a matrix element. One of those MVAs matched the jets under the hypothesis that the event is a signal  $tHq$  event and the second one matches the jets under the hypothesis that the event is a  $t\bar{t}$  event. Reconstructing all events under these two hypothesis allows the construction of variables separating signal especially well against the dominant  $t\bar{t}$  background.

The higher the multiplicity of jets in an event is, the more possibilities arise to assign jets to the quarks of the specific process. Therefore we apply constraints reducing the number of jet assignment possibilities.

## 2.4. Event Classification

Events are then classified as either signal or background by employing a third MVA, which uses three categories of input variables. The first category consists of observables which are not dependent on any of the jet assignments done before and exploit global properties of an event. The second category exploits observables obtained with the  $tHq$  reconstruction and the third category uses observables obtained with the  $t\bar{t}$  reconstruction. A total number of eight input variables are used to separate signal from background. In Figure 2 the eight different input variables can be seen sorted into the three different categories. As an example the most discriminating variable, the absolute value of the pseudorapidity of the reconstructed light jet, is shown in Figure 3. The response distributions of the classification MVA for signal and background events can be seen in Figure 4. A good separation can be achieved.

Category	Input Variable
global variables	Electric charge of the the lepton
tHq reconstruction	Pseudorapidity of the recoil jet
	Number of b-tagged jets among the two jets from the Higgs boson decay
	Transverse momentum of the Higgs boson
	Transverse momentum of the recoil jet
t $\bar{t}$ reconstruction	$\Delta R$ between the two light-flavor jets from the decay of $t_{had}$
	Mass of $t_{had}$
	Number of b-tagged jets among the two light-flavor jets from the decay of $t_{had}$

Figure 2: Input variables for the classification MVA. The variables are split into three groups: global variables, variables of the jet assignment under the  $tHq$  hypotheses, variables of the jet assignment under the  $t\bar{t}$  hypothesis. In the descriptions,  $t_{had}$  stands for a hadronically decaying top quark.

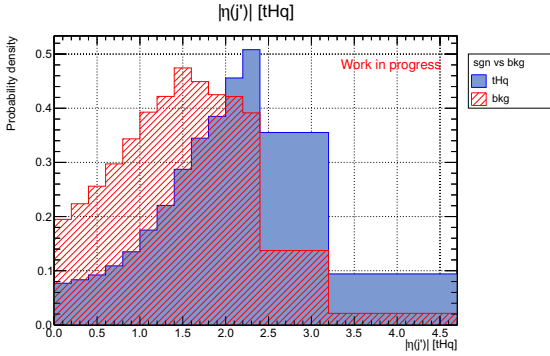


Figure 3: Distributions of signal and background events over the absolute value of the pseudorapidity of the reconstructed light jet, which is used as input variable for the classification MVA. This variable is the single most discriminating variable between signal and background events.

### 3. Summary

With the full dataset at a center of mass energy of 8 TeV a search for t-channel single top production in association with the 125 GeV Higgs boson is conducted. The tHq decay channel is used with  $H \rightarrow b\bar{b}$ . The signal is assumed to come from the  $y_t = -1$  coupling which would require beyond the standard model physics. The assignment of the jets is using two separate artificial neural networks. This analysis pioneers the technique, that each event is reconstructed under a potential tHq signal hypothesis and a  $t\bar{t}$  background hypothesis. The individual reconstructions provide input variables with a high separation power between signal and background events. A third neural network then uses these input variables, obtained from the reconstruc-

tions, to enhance the signal-to-background ratio and discriminate against the main backgrounds, including the predominant  $t\bar{t}$  production. This process is modelled with a simulation-driven approach and a data-driven approach cross-checking each other. The analysis was not yet public at the time of the conference, but results have been made public shortly afterwards [4].

### References

- [1] S. Chatrchyan, et al., Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, J. High Energy Phys. 06 (2013) 081. doi:10.1007/JHEP06(2013)081.
- [2] G. Aad, et al., Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys.Lett. B716 (2012) 1–29. arXiv:1207.7214, doi:10.1016/j.physletb.2012.08.020.
- [3] M. Farina, F. Grojean, F. Maltoni, E. Salvioni, A. Thamm, Lifting degeneracies in higgs couplings using single top production in association with a higgs boson, JHEP 1305 (2013) 022. arXiv:1211.3736, doi:10.1007/JHEP05(2013)022.
- [4] S. Chatrchyan, et al., Search for H to  $b\bar{b}b\bar{b}$  in association with single top quarks as a test of Higgs couplings (CMS-PAS-HIG-14-015).

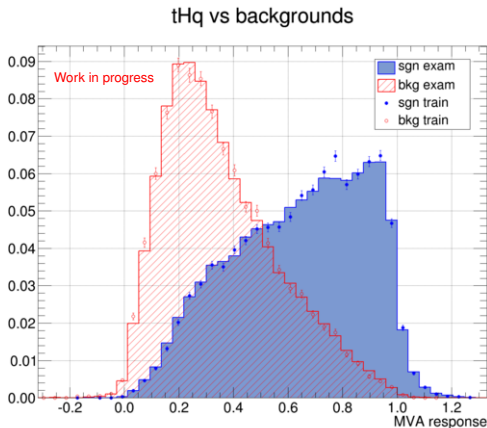


Figure 4: Distributions of the classification MVAs response for signal and background events. A good separation has been achieved.